

Adding a referee to an interconnection network: What can(not) be computed in one round. *

F. Becker¹, M. Matamala^{2,3}, N. Nisse⁴, I. Rapaport^{2,3}, K. Suchan^{5,6}, I. Todinca¹

¹LIFO, Université d'Orléans, France.

²Departamento de Ingeniería Matemática, Universidad de Chile, Santiago, Chile.

³Centro de Modelamiento Matemático (UMI 2807 CNRS), Univ. de Chile, Santiago, Chile.

⁴MASCOTTE, INRIA, I3S (CNRS/Univ. Nice Sophia Antipolis), France.

⁵Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Santiago, Chile.

⁶Faculty of Applied Mathematics, AGH - Univ. of Science and Technology, Cracow, Poland.

Abstract

In this paper we ask which properties of a distributed network can be computed from a few amount of local information provided by its nodes. The distributed model we consider is a restriction of the classical *CONGEST* (distributed) model and it is close to the simultaneous messages (communication complexity) model defined by Babai, Kimmel and Lokam. More precisely, each of these n nodes -which only knows its own ID and the IDs of its neighbors- is allowed to send a message of $O(\log n)$ bits to some central entity, called the referee. Is it possible for the referee to decide some basic structural properties of the network topology G ? We show that simple questions like, "does G contain a square?", "does G contain a triangle?" or "Is the diameter of G at most 3?" cannot be solved in general. On the other hand, the referee can decode the messages in order to have full knowledge of G when G belongs to many graph classes such as planar graphs, bounded treewidth graphs and, more generally, bounded degeneracy graphs. We leave open questions related to the connectivity of arbitrary graphs.

1 Introduction

When referring to a "network", one typically visualizes it as a distributed system where nodes correspond to agents or processors. These nodes can only interact locally and therefore, since they all lack of global information, new algorithmic and complexity notions arise. In contrast with classical algorithmic theory – where the Turing machine is the consensus formal model of algorithm – in the case of distributed systems there are many models for communication protocols.

Some theoretical and over-simplified models were conceived for studying particular aspects of protocols such as fault-tolerance, synchronism, locality, congestion, etc. In the model *CONGEST* (see the book of Peleg [16]) a network is represented by a graph whose nodes correspond to network processors and edges to inter-processor links. The communication is synchronous and occurs in

*Partially supported by programs Fondap and Basal-CMM (M.M., I.R., K.S.), Fondecyt 1090156 (I.R.), Fondecyt 1100192 (M.M.), Fondecyt 11090390 and Anillo ACT88 (K.S.), Instituto Milenio ICDB (I.R.), ECOS-Conicyt C09E04 (M.M., I.R., I.T.), ANR blanc AGAPE and ALADDIN (I.T.)

discrete rounds. In each round every processor can send a message of size $O(\log n)$ through each of its outgoing links (where n is the number of nodes).

Some variations to the *CONGEST* model have been proposed. The general idea is to remove some restrictions (making it more powerful) in order to focus on some particular issues. In that spirit Linial [12] introduced in a seminal paper the free model (also called *LOCAL* [16]). The only difference with the *CONGEST* model is that the restriction on the size of the messages is removed so that every vertex is allowed to send unbounded size messages in every round. By relaxing the constraint on the size of the messages, this model focuses on the issue of locality in distributed systems. For that reason, the answer to the question *What cannot be computed locally?* [11] (in the *LOCAL* model) appears to be a crucial one. In this context, Kuhn *et al.* show that difficult problems like minimum vertex cover and minimum dominating set cannot be well approximated when processors can exchange arbitrary long messages during a bounded number of rounds [11].

Grumbach and Wu [8] also use the *CONGEST* model. They are interested in *frugal computations*, i.e., computations where the total amount of information traversing each edge is $O(\log n)$. Their motivation was to understand the impact of locality. In fact, they show that for planar networks or networks of bounded degree, any first order logic formula can be evaluated frugally (in their setting).

We propose in this paper an alternative model for frugal computation. In our model, the number of communication rounds is bounded, but, at each round, every vertex can send (and receive) a message of size $O(\log n)$ to (from) a central entity, called *the referee* which communicates with all nodes of the network. The notion of referee has been used, e.g., in the *SIMULTANEOUS MESSAGES* model of [2]. We have the same motivation of Grumbach and Wu [8]. More precisely, we want to investigate the computational power of our distributed model by deciding network topology properties¹.

1.1 Our results

The positive result of this work says that, given $k \in \mathbb{N}$, there exists a *one-round* protocol allowing the referee to reconstruct graphs of degeneracy k (or to decide that the graph degeneracy is bigger than k , see Section 3 for definitions). Note that, clearly, n -node graphs with bounded degeneracy can be encoded using $O(n \log n)$ bits. The question here is to know whether this amount of *local* information is sufficient to encode such graphs. Notice also that forests have degeneracy 1 while planar graphs have degeneracy 5 and that the degeneracy of a graph is upper bounded by its treewidth. So our protocol performs well in all these graph classes.

On the negative side, we prove that there is no one-round protocol allowing the referee to decide whether an arbitrary graph G contains a square, resp., a triangle, nor for computing its diameter.

Two central issues remain open. On the one hand, the existence of a one-round-frugal protocol to decide connectivity. On the other hand, can we decide more properties by allowing more rounds?

1.2 The model

An interconnection network is modeled by a simple undirected connected n -node graph $G = (V, E)$. Each node $x \in V$ has a unique identifier $ID(x)$ between 1 and n . In particular, in the whole paper,

¹On some aspects, for instance on graphs of bounded degree, our model is much more powerful than theirs. It is clear, that if the network has bounded degree then each processor can simply send its neighborhood to the referee, using only $O(\log(n))$ bits. And, with this information, the referee is able to *reconstruct* the whole network.

"graph" means "labelled graph". In the sequence of vertices denoted by (v_1, \dots, v_n) , v_i denotes the vertex $x \in V$ such that $ID(x) = i$. At each node v there is a local processing unit that knows its identifier, the local set $\{ID(y) \mid y \in N_G(v)\}$ of identifiers of the neighbors of v and the total number of nodes n , where $N_G(v)$ is the set of neighbors of node v .

We consider the particular case where there is a central entity, the referee. That is, the interconnection network \mathcal{G} consists of the union of a graph $G = (V, E)$, with $V = \{v_1, \dots, v_n\}$, plus a universal node v_0 representing the referee, i.e., v_0 is adjacent to any node of G .

At each round of the communication process, any node may perform a local computation based on its own knowledge and then send and receive one message to (from) each of its neighbors. The messages may be different for any neighbor. In particular, at each round, the referee may receive one message from any node in G . The protocol is said *frugal* if the size of each message is limited to $O(\log n)$ bits. We distinguish the communication time complexity that corresponds to the number of rounds of the communication process and the local time complexity corresponding to the maximum time taken by the local computations.

In this paper, we study one-round frugal protocol for solving problems on the graph G through the interconnection network \mathcal{G} . In other words, the referee, having no knowledge of the network topology but its size, must solve some problems using the n messages, each of size $O(\log n)$ bits, received from the n processors. Note that, since we only consider a single round of communication, the network may be asynchronous. Indeed, the referee can wait until it has received one message from every vertex (this only requires that the referee knows the size of the network).

A one-round protocol is thus made of a local phase, where each node computes a message as a function of its neighborhood, and a global phase, where the *referee* computes the result of the computation from the messages.

Definition 1 A one-round protocol Γ is a family $(\Gamma_n^l, \Gamma_n^g)_{n \in \mathbb{N}}$, where:

- $\Gamma_n^l : \{1, \dots, n\} \times \mathcal{P}(\{1, \dots, n\}) \rightarrow \{0, 1\}^*$ is the local function of Γ for graphs of size n ,
- $\Gamma_n^g : (\{0, 1\}^*)^n \rightarrow \{0, 1\}^*$ is the global function of Γ for graphs of size n .

Given $G = (V = \{1, \dots, n\}, E)$, the message vector of Γ on G is:

$$\Gamma^l(G) = (\Gamma_n^l(1, N_G(1)), \dots, \Gamma_n^l(n, N_G(n))).$$

The output of Γ on G is:

$$\Gamma(G) = \Gamma_n^g(\Gamma^l(G)).$$

We define

$$|\Gamma^l(G)| = \max_{1 \leq i \leq n} |\Gamma_n^l(i, N_G(i))|.$$

Γ is said to be frugal if:

$$\max_{G \text{ graph of } n \text{ nodes}} (|\Gamma^l(G)|) = O(\log n).$$

Note that we do not care about the complexity (or computability, or uniformity) of Γ_n^l and Γ_n^g , in agreement with the usual setting of communication complexity.

Notice that function Γ_n^l can be evaluated in any pair (i, N) where $i \in \{1, \dots, n\}$ and $N \subseteq \{1, \dots, n\}$. Then, $\Gamma_n^l(i, N)$ corresponds to the message sent to the referee by node i in a graph of n nodes when its neighborhood is N .

1.3 Related Work

Testing properties of graphs has been first investigated (in a centralized point of view) by Goldreich *et al.* [6] (see [5] for a survey). Given the adjacency matrix of a graph, the question is to determine the minimum number of elementary queries (e.g., "what is the i^{th} neighbor of some vertex v ?") necessary to decide whether the graph satisfies some given property. In this context, probabilistic algorithms are given that always accept a graph if it satisfies the property and reject with constant probability any graph that is "far enough" from the property. For instance, [1] gives lower and upper bounds for testing the triangle freeness in general graphs. Closer to our model, in [7], Goldreich and Ron provide efficient such algorithms for testing the connectivity of bounded degree graphs when they are given the list of neighbors of every vertex.

Other tradeoffs between the size of a data structure and the complexity (in terms of number of bits that must be checked in this structure) of algorithms for answering some queries have been provided using the computational complexity model and the cell probe model [21, 22, 13, 14]. Testing graph properties has also been widely investigated using these frameworks (e.g., [19, 3, 15]).

This paper is organized as follows. In Section 2 we present our negative results: there is no frugal protocol deciding if the graph is of small diameter, or if it contains a square or a triangle. In Section 3 we provide a frugal protocol reconstructing graphs of bounded degeneracy. We conclude with some open questions in Section 4.

2 Hard local properties

It turns out that given a small non-trivial graph S , the question: *does G admit S as a (not necessarily induced) subgraph?* is most often impossible to answer in one round. One reason for this is that if S is not reduced to an edge, then none of the vertices of G knows which of their neighbors are susceptible to be involved in an instance of S : to them, they all look the same! Because of this, they would need to send their whole adjacency list to the referee. This, of course, represents $O(n \log n)$ bits, while each vertex is allowed to send only $O(\log n)$ bits.

In this section we show an example of such behavior, where S is a square. To obtain our impossibility results we use two ingredients. We first prove in Lemma 1 that if a family of graphs can be reconstructed by one-round, frugal protocol, then the number of graphs of size n , in the family, is $2^{O(n \log n)}$. Then, we introduce a reduction technique: given a one-round, frugal protocol for deciding whether S is a subgraph of G , we will build a one-round, frugal protocol for reconstructing a certain family of graphs. When S is a square, this family is the class of graphs without squares. We get the impossibility result as the number of graphs without squares (as subgraphs) of size n is $2^{\Theta(n^{3/2})}$ [9].

Using the same technique, we show that one cannot determine whether a graph has a diameter less than four. In that case, the reduction goes from computing the diameter of a family of auxiliary graphs to reconstructing the original graph. In other words, the family we use there is the family of all graphs.

Lemma 1 *Let \mathcal{G} be a family of graphs, and $g(n)$ be the number of graphs in \mathcal{G} with set of vertices $\{1, \dots, n\}$. If there is a frugal one-round protocol for reconstructing graphs in \mathcal{G} (i. e. a protocol whose output on any $G \in \mathcal{G}$ is its adjacency matrix), then $\log g(n) = O(n \log n)$.*

Proof Let Γ be a frugal one-round algorithm. There is a k such that the referee receives at

most $k \cdot \log n$ bits per vertex on a graph with n vertices, hence $kn \log n$ bits in total. From this information Γ must reconstruct $g(n)$ different graphs. Hence, $\log g(n) \leq kn \log n$. \square

2.1 Finding a square in one round is hard

Theorem 1 *There is no one-round frugal protocol allowing the referee to decide whether an arbitrary graph G contains a square as a subgraph.*

Proof For the sake of contradiction, we assume that a one-round frugal protocol Γ exists for deciding whether a graph admits the square as a subgraph, and show how to derive a frugal one-round protocol Δ for reconstructing graphs without squares. This would contradict Lemma 1 since there are $2^{\Theta(n^{3/2})}$ graphs without squares and with n vertices [9].

Let G be a graph without squares with n vertices. For each $s, t \in \{1, \dots, n\}$, $s \neq t$, we are going to simulate the behavior of Γ on the graph $G'_{s,t}$, obtained from G by adding n new vertices, numbered from $n+1$ to $2n$, n new edges, $\{i, n+i\}$, for $i \in \{1, \dots, n\}$, and the edge $\{n+s, n+t\}$. Graph $G'_{s,t}$ contains a square if and only if s and t are adjacent in G .

Note that, for every value of s and t , the neighborhood of $i \in \{1, \dots, n\}$ in the graph $G'_{s,t}$ is exactly the same: $N_G(i) \cup \{i+n\}$. Our protocol Δ works as follows. Each vertex $i \in \{1, \dots, n\}$ constructs and sends exactly the same message as protocol Γ , on vertex i , in the graph $G'_{s,t}$. The referee collects all these messages and simulates, for each couple of values s, t , the messages that protocol Γ would send for graph $G'_{s,t}$, for each vertex $j \in \{n+1, \dots, 2n\}$ (these messages do not depend on G). Then, applying the global function of Γ , it decides if $G'_{s,t}$ has a square, i.e. it decides if s and t are adjacent in G ; thus graph G can be reconstructed. \square

Algorithm 1 GlobalFunction Δ_n^g – ReconstructGraphsWithoutSquares

Ensure: $H = (V, E)$ – a reconstruction of G

Require: $\forall i \in \{1, \dots, n\}$, the message m_i , sent by the node i of G and computed using the local function Δ_n^l defined as: $m_i = \Delta_n^l(i, N_G(i)) = \Gamma_{2n}^l(i, N_G(i) \cup \{i+n\})$

for all $s, t \in \{1, \dots, n\}$, $s \neq t$ **do**

for all $j \in \{n+1, \dots, 2n\} \setminus \{n+s, n+t\}$ **do**

$m_j(s, t) \leftarrow \Gamma_{2n}^l(j, \{j-n\})$ – This does not depend on G

end for

$m_{n+s}(s, t) \leftarrow \Gamma_{2n}^l(n+s, \{s, n+t\})$

$m_{n+t}(s, t) \leftarrow \Gamma_{2n}^l(n+t, \{t, n+s\})$

if $\Gamma_{2n}^g(m_1, \dots, m_n, m_{n+1}(s, t), \dots, m_{2n}(s, t)) = 1$ **then**

$\{s, t\} \in E$ – Since $G'_{s,t}$ has a square

else

$\{s, t\} \notin E$ – Since $G'_{s,t}$ has no square

end if

end for

return $H = (\{1, \dots, n\}, E)$, as reconstructed above

By the same arguments we deduce that there is no frugal one-round protocol testing if the graph has a square as an *induced* subgraph.

2.2 Computing diameter in one round is hard

In the case of squares, proving the impossibility of a one-round frugal protocol turned out to be easy because we could construct, from a graph G , a family of graphs $G'_{s,t}$ where the neighborhoods of the original vertices of G in $G'_{s,t}$ did not depend on s and t . We will use a variant of that trick to prove that it is not possible to decide whether a graph has a small (≤ 3) diameter. Again, for each couple of vertices s and t of the original graph, we consider a new graph $G'_{s,t}$, but this time the neighborhoods of the original vertices may depend on s and t . Nevertheless, each original vertex will have only three possible neighborhoods in the family of graphs $G'_{s,t}$, which allows us to keep the reduction technique.

Theorem 2 *There is no one-round protocol allowing the referee to decide whether the diameter of an arbitrary graph is at most 3.*

Proof Once again, for the sake of the contradiction, let us assume that there is a frugal one-round protocol Γ for deciding if the diameter of the graph is at most 3. We shall prove that by using Γ we can build a frugal one-round protocol Δ for reconstructing any graph. Since there are $\Omega(2^{n^2/2})$ graphs with vertices $\{1 \dots n\}$, we would get a contradiction with Lemma 1.

Informally, for each $s, t \in \{1, \dots, n\}$, we shall simulate the behavior of Γ on the graph $G'_{s,t}$ (Figure 1) obtained from a graph $G = (V, E)$ by adding three extra vertices $n+1$, $n+2$ and $n+3$, and edges $\{s, n+1\}$, $\{t, n+2\}$, and $\{v, n+3\}$ for all $v \in \{1, \dots, n\}$. $G'_{s,t}$ has a diameter ≤ 3 if and only if $\{s, t\} \in E$: all vertices of index $\leq n$ are at distance at most 2, by going through $n+3$, so only $n+1$ and $n+2$ can be at distance 4, which happens if $\{s, t\} \notin E$. That way, since for each s, t , protocol Γ can decide whether $G'_{s,t}$ is of diameter at most 3, it will be possible to reconstruct G .

On an input graph $G = (V, E)$, the local function of Δ computes $m_i^0 := \Gamma_{n+3}^l(i, N_G(i) \cup \{n+3\})$, $m_i^s := \Gamma_{n+3}^l(i, N_G(i) \cup \{n+1, n+3\})$ and $m_i^t := \Gamma_{n+3}^l(i, N_G(i) \cup \{n+2, n+3\})$, for each $i \in \{1, \dots, n\}$ and the message sent to the referee is the triple (m_i^0, m_i^s, m_i^t) . Thus, Δ is frugal, since its messages are three times as big as those of Γ .

The global function of Δ works as follows. For each $s, t, i \in \{1, 2, \dots, n\}$, the referee computes $m_i(s, t)$, defined as m_i^s if $i = s$, as m_i^t if $i = t$ and as m_i^0 if $i \notin \{s, t\}$. For $i \in \{n+1, n+2, n+3\}$, $m_i(s, t)$ is defined as the Γ_{n+3}^l function on vertex i in $G'_{s,t}$, and can be computed by the referee since it does not depend on G , but only on Γ , s and t . Note that, for any $i \leq n+3$, message $m_i(s, t)$ is precisely the message that the local function Γ_{n+3}^l would send for vertex i in the graph $G'_{s,t}$. Consequently, $\Gamma_{n+3}^g(m_1(s, t), \dots, m_{n+3}(s, t))$ accepts if and only if s and t are adjacent in graph G . Thus, from the messages m_i^0 , m_i^s and m_i^t , protocol Δ can construct messages $m_i(s, t)$ and eventually reconstruct the whole graph G .

We define Δ_n^g formally in algorithm 2. □

2.3 Finding a triangle in one round is hard

Using similar ingredients as in Theorem 2, we can prove that one can not frugally detect the existence of a triangle.

Theorem 3 *There is no one-round protocol allowing the referee to decide whether an arbitrary graph G contains a triangle as a subgraph.*

Algorithm 2 GlobalFunction Δ_n^g – ReconstructGraph

Ensure: $H = (V, E)$ – a reconstruction of G

Require: $\forall i, (m_i^0, m_i^s, m_i^t)$ the messages sent by the local function, where

$$m_i^0 = \Gamma_{n+3}^l(i, N_G(i) \cup \{n+3\})$$

$$m_i^s = \Gamma_{n+3}^l(i, N_G(i) \cup \{n+1, n+3\})$$

$$m_i^t = \Gamma_{n+3}^l(i, N_G(i) \cup \{n+2, n+3\})$$

for all $s, t \in \{1, \dots, n\}$ **do**

$$m_s \leftarrow m_s^s \text{ // encoding of the neighborhood of } s \text{ in } G'_{s,t}$$

$$m_t \leftarrow m_t^t \text{ // encoding of the neighborhood of } t \text{ in } G'_{s,t}$$

for all $i \notin \{s, t\}$ **do**

$$m_i \leftarrow m_i^0 \text{ // encoding of the neighborhood of } i \text{ in } G'_{s,t}$$

end for

$$m_{n+1} \leftarrow \Gamma_{n+3}^l(n+1, \{s\})$$

$$m_{n+2} \leftarrow \Gamma_{n+3}^l(n+2, \{t\})$$

$$m_{n+3} \leftarrow \Gamma_{n+3}^l(n+3, \{1, \dots, n\})$$

if $\Gamma_{n+1}^g(m_1, \dots, m_{n+3}) = 1$ **then**

$$\{s, t\} \in E \text{ // Since } G'_{s,t} \text{ has a diameter } \leq 3$$

else

$$\{s, t\} \notin E \text{ // Since } G'_{s,t} \text{ has a diameter } > 3$$

end if

end for

return $H = (\{1, \dots, n\}, E)$, as reconstructed above

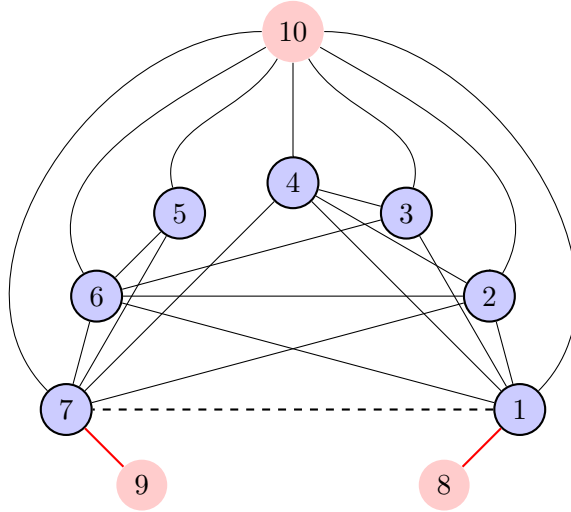


Figure 1: Reducing reconstruction of arbitrary graphs to computation of the diameter: given the graph G (circled vertices), in order to check whether $(1, 7)$ is an edge of G , we build the auxiliary graph $G'_{1,7}$ by adding vertices 8 to 10 as depicted on the figure. It has diameter 3 iff $(1, 7)$ is an edge of G (and thus of $G'_{1,7}$), otherwise it has diameter 4: in both cases, the longest path goes from 8 to 9.

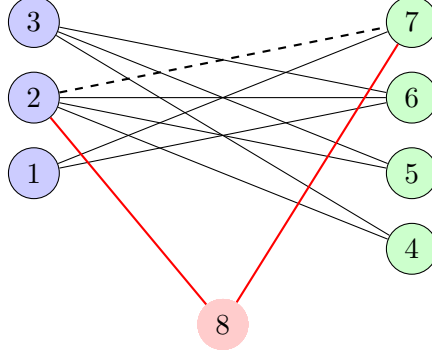


Figure 2: Reducing reconstruction of arbitrary graph to detection of triangles: given the graph G (circled vertices), in order to check whether $(2, 7)$ is an edge of G , we build the auxiliary graph $G'_{2,7}$ by adding vertex 8 as depicted on the figure. It contains a triangle iff $(2, 7)$ is an edge of G .

Proof Once again, for the sake of the contradiction, let us assume that there is a frugal one-round protocol Γ for detecting triangles. We shall prove that by using Γ we can build a frugal one-round protocol Δ for reconstructing bipartite graphs with parts $\{1, \dots, n/2\}$ and $\{n/2 + 1, \dots, n\}$. Since there are $\Omega(2^{(n/2)^2})$ such bipartite graphs we would get a contradiction with Lemma 1.

Informally, for each $s, t \in \{1, \dots, n\}$, we shall simulate the behavior of Γ on the graph $G'_{s,t}$ obtained from a graph $G = (V, E)$ by adding an extra vertex $n + 1$, and edges $\{s, n + 1\}$ and $\{t, n + 1\}$. If G is bipartite as above, then $G'_{s,t}$ has a triangle if and only if $\{s, t\} \in E$ (see Figure 2). Therefore, since for each s, t , protocol Γ can decide whether $G'_{s,t}$ has a triangle, it will be possible to reconstruct G .

On an input graph $G = (V, E)$, the local function of Δ computes $m'_i := \Gamma_{n+1}^l(i, N_G(i))$ and $m''_i := \Gamma_{n+1}^l(i, N_G(i) \cup \{n + 1\})$, for each $i \in \{1, \dots, n\}$ and the message sent to the referee is the couple (m'_i, m''_i) . Thus, Δ is frugal, since its messages are twice as big as those of Γ .

Let us describe the global function of Δ . For each s, t , and for $i \in \{1, \dots, n + 1\}$, let $m_i(s, t)$ be defined as m'_i when $i \notin \{s, t, n + 1\}$, as m''_i when $i \in \{s, t\}$, and as $\Gamma_{n+1}^l(n + 1, \{s, t\})$ when $i = n + 1$. Then, for each $i \in \{1, \dots, n + 1\}$, $m_i(s, t)$ is the message that the local function Γ_{n+1}^l would send to the referee when evaluated in node i of the graph $G'_{s,t}$. Therefore, $\Gamma_{n+1}^g(m_1(s, t), \dots, m_{n+1}(s, t))$ accepts if and only if $\{s, t\}$ is an edge of G . Thus, from the messages m'_i and m''_i , protocol Δ can construct messages $m_i(s, t)$ and eventually reconstruct the whole graph G . \square

It is worth mentioning that in Theorems 1, 2 and 3, we provide reductions showing that if there exists a one-round protocol detecting squares (resp., triangles, resp., long distances) and using messages of $k(n)$ bits per node (on graphs on n vertices), then there exist one-round protocols reconstructing graphs without squares (resp. any graph with the given partition) using $k(2n)$ (resp. $3k(n + 3)$) bits. Moreover, the local time complexities of new protocols are polynomially bounded in terms of the original protocols.

3 Reconstruction of graphs of bounded degeneracy

We say that G is of degeneracy k if there is a vertex r of degree at most k in G that we can remove, and then proceed recursively on the resulting graph $G' = G \setminus r$, until we obtain an empty graph.

Let us denote by r_i the i -th vertex removed from the graph.

Definition 2 $G = (V, E)$ is of degeneracy k if there exists a permutation (r_1, \dots, r_n) of V such that for all i , $1 \leq i \leq n$, r_i is of degree at most k in G_i , where G_i is the subgraph of G induced by $\{r_1, \dots, r_i\}$.

In this section, we show that, when the underlying graph G is of bounded degeneracy, there is a one-round frugal protocol reconstructing the graph.

3.1 Case of forests ($k = 1$)

To give the flavour of the algorithm for bounded degeneracy graphs, we start with the graphs with degeneracy 1, which is exactly the class of forests. Let T be a forest, and let $N_S(v)$ denote the neighborhood of $v \in V(S)$ in the subforest S of T .

In this case, any vertex v sends a triple of integers to the referee. This triple consists of its identifier, its degree in T , $\deg_T(v)$, and the sum of the identifiers of its neighbors (this clearly can be encoded using less than $4 \log n$ bits):

$$(ID(v), \deg_T(v), \sum_{w \in N_T(v)} ID(w)).$$

To decode the information, the referee chooses a leaf v (one of the vertices with degree at most 1). Intuitively, the referee prunes this leaf from the forest. By doing this recursively, it gets all the information concerning the forest. More precisely, the triple of v contains the identifier of the unique neighbor w of v since the sum of the IDs of the neighbors of v is exactly $ID(w)$. The referee can replace the triple of w by

$$(ID(w), \deg_T(w) - 1, (\sum_{z \in N_T(w)} ID(z)) - ID(v))$$

which is exactly

$$(ID(w), \deg_{T \setminus v}(w), \sum_{z \in N_{T \setminus v}(w)} ID(z)).$$

Combining this new triple with all triples of the vertices but v allows to have the information on $T \setminus v$. By induction on the number of vertices, the referee is able to decode this information and rebuild the whole forest (or decide whether the graph contains a cycle).

In the following, we generalize the idea of "pruning" a vertex v (of degree at most k) from the graph G in such a way that the information of the pruned graph $G \setminus v$ can be obtained from the information of all vertices of G (by modifying the information of the neighbors of v).

3.2 Generalization for any $k \geq 1$

Each vertex needs to know the value of k (recall that G is of degeneracy at most k). Moreover, some data structures that allow working on graphs of degeneracy upper bounded by k have to be present at all vertices. Nevertheless, no elimination order (see Definition 2) can be known a priori: it will be discovered during the execution of calculations by the referee.

The information $\Gamma_n^l(v, N_G(v))$ that each node v sends to the referee is the following $k + 2$ -tuple:

- its identifier $ID(v)$.
- its degree $deg(v)$ in G .
- for each integer p , $1 \leq p \leq k$, the quantity $\sum_{w \in N_G(v)} (ID(w))^p$ (i.e., the sum of p 's powers of the identifiers of the neighbours).

Note that for $k = 1$ this is the construction described in the case of forests. We shall see that, with this information, for any vertex v of degree at most k , the universal vertex can retrieve the identifiers of the neighbours of v . Eventually, like in the case of forests, the referee simulates the removal of node v from the graph and iterates the process until obtaining the empty graph.

To describe the encoding and decoding of the neighborhood information we need to recall some results from algebra and number theory. We will use the following matrix, very similar to the well-known Vandermonde matrix.

Definition 3 Define the matrix A by $A_{p,i} = i^p$, for $i = 1, \dots, n$ and $p = 1, \dots, k$. To express explicitly the dimensions we will write $A(k, n)$.

3.3 Neighborhood encoding

Let us recall that (v_1, \dots, v_n) denotes the vertices ordered by their identifiers. To encode its neighborhood, each vertex x uses the matrix $A(k, n)$ and the incidence vector of its neighborhood \mathbf{x} , i.e., the binary vector with 1 on the i -th coordinate if v_i is a neighbor of x , and 0 otherwise.

Algorithm 3 LocalFunction

Require: $k, n, A(k, n), x, deg(x), \mathbf{x}$, where \mathbf{x} is the incidence vector of the neighborhood of x

Ensure: $\mathbf{b}(x) = A(k, n)\mathbf{x}$

$\mathbf{b}(x) = A(k, n)\mathbf{x}$

return $(ID(x), deg(x), \mathbf{b}(x))$

Lemma 2 The size of the message generated in Algorithm 3 is $O(\log n)$ bits – more precisely, $O(k^2 \log n)$ bits. The computation can be performed in $O(n)$ local time.

Proof For computing \mathbf{b} , the algorithm sums up at most n columns of the matrix $A(k, n)$. The result is a k element vector with a sum of some elements from i -th row of $A(k, n)$ at position i . The coefficients in $A(k, n)$ are at most n^k , so the sum is at most n^{k+1} . It can be encoded on $(k+1) \log n$ bits, so the whole vector \mathbf{b} takes $k(k+1) \log n$ bits at most. Altogether, the message associated to v is of size at most $O(k^2 \log n)$.

For the time complexity it is enough to notice that we sum up $O(n)$ values encoded on $O(\log n)$ bits each. \square

3.4 Neighborhood decoding

We will use the following classical result from number theory.

Theorem 4 [20] *In integers, the following system of simultaneous equations has no non-trivial solutions, i.e., it implies that $(i_1, i_2, \dots, i_k) = (j_1, j_2, \dots, j_k)$ up to a permutation.*

$$i_1^p + i_2^p + \dots + i_k^p = j_1^p + j_2^p + \dots + j_k^p \text{ for all } 1 \leq p \leq k$$

Notice that Theorem 4 covers also the case where some variables are equal 0, so we have the following corollary.

Corollary 1 *Given $(ID(x), deg(x), \mathbf{b}(x))$ obtained by Algorithm 3 for a vertex of degree at most k , there exists only one binary vector \mathbf{x} such that $A(k, n)\mathbf{x} = \mathbf{b}(x)$.*

In case a fast decoding of neighborhoods is needed, we can perform a preprocessing step to enumerate all k -subsets of $\{1, \dots, n\}$ and compute the values $\mathbf{b} = A(k, n)\mathbf{x}$, where \mathbf{x} is an incidence vector of such a subset, and store them in a table N that assigns to each value vector \mathbf{b} the corresponding \mathbf{x} . The size of N is $O(n^k)$ and, by sorting it according to the lexicographic order on value vectors, we can perform a neighborhood look-up in time $k \log n$. Thus we have the following.

Lemma 3 *Let k, n be integers. There exists a function that, given $(ID(x), deg(x), \mathbf{b}(x))$ generated by Algorithm 3 for a vertex x of degree at most k , allows to compute the neighborhood of x in time $O(\log n)$.*

Using such a look-up table we can perform Algorithm 4, which reconstructs graph G in $O(n^2)$ time.

Algorithm 4 GlobalFunction

Require: $\mathcal{B} = \{(ID(x), deg(x), \mathbf{b}(x)) \mid x \in V\}$, look-up table N (as in Lemma 3)

Ensure: $H = (V, E)$ – a reconstruction of G

while there is an element in \mathcal{B} **do**

 take an element $(ID(x), deg(x), \mathbf{b}(x))$ from \mathcal{B} s.t. $deg(x) \leq k$

 look-up in N the neighborhood \mathbf{x} of x

for all $v_i \in V$ s.t. $\mathbf{x}(i) = 1$ **do**

 add $\{x, v_i\}$ to H

 update $(ID(v_i), deg(v_i), \mathbf{b}(v_i))$ in \mathcal{B} according to the removal of x from G . That is, $deg(v_i)$ is decreased by one and, for any $1 \leq p \leq k$, the p^{th} coordinate $\mathbf{b}_p(v_i)$ of $\mathbf{b}(v_i)$ is replaced by $\mathbf{b}_p(v_i) - ID(x)^p$

end for

 remove $(ID(x), deg(x), \mathbf{b}(x))$ from \mathcal{B}

end while

return H

We deduce:

Theorem 5 *There exists a one-round frugal protocol allowing the referee to reconstruct graphs of bounded degeneracy.*

Note that our protocol can also be turned into a recognition protocol for these graphs. By applying the same encoding and decoding algorithm as above, we just have to add one test in

Algorithm 4, which rejects the graph if, during the pruning process, we find no vertex of degree at most k .

Many graph classes are known to be of bounded degeneracy. Planar graphs are of degeneracy at most 5, graphs of treewidth k are also of degeneracy at most k , and more generally, for each fixed graph H , the class of H -minor free graphs is also of bounded degeneracy [10, 17, 18]. Eventually, we point out that we can also work with complements of neighborhoods. More precisely, we can define graphs of "generalized degeneracy k " as the graphs $G = (V, E)$ having a vertex ordering (r_1, \dots, r_n) of V such that for all i , $1 \leq i \leq n$, r_i is of degree at most k in G_i or in the complement of G_i (as in Definition 2, G_i is the subgraph of G induced by $\{r_1, \dots, r_i\}$). We can adapt our protocol for the reconstruction of graphs of generalized degeneracy at most k , by encoding both the neighborhood and the non-neighborhood of each vertex.

4 Conclusion and open questions

We have investigated a model of distributed computing combining constraints of locality and congestion. Somehow surprisingly, we have seen that the model is powerful enough to reconstruct "complicated" topologies (graphs of bounded degeneracy, which include e.g. planar graphs), while "simple" properties, like deciding if a graph has a triangle or if a graph has diameter at most three cannot be decided in the model. We point out that the hardness results concern both local properties (triangles) and global properties (diameter).

The main open question is the existence of a one-round frugal protocol deciding if a graph is connected. We rather tend to believe there is no such protocol. Nevertheless, our techniques for hardness results fail in this case because they are based on a partitioning argument with a fixed number of parts. In our hardness proofs, we have partitioned the vertices of the graph into two or three parts, and we have shown that, even if vertices of a same part are allowed to share their local information, the problem remains intractable. Such arguments cannot work in the case of connectivity: if a graph is split into k parts and vertices of each part are allowed to communicate to each other, there is an algorithm for connectivity using $O(k \log n)$ bits per node. Therefore we need to invent different techniques. We point out that this type of difficulty for showing lower bounds also arises in classical multiparty communication complexity [4].

Another natural question is whether one can find a frugal one-round protocol deciding if a graph is bipartite. As ongoing work, we have proved that the existence of a frugal one-round protocol for bipartiteness implies the existence of a frugal one-round protocol deciding if a bipartite graph is connected.

We think that our reduction technique is a promising tool for future research, in particular we believe it is worth to study the existence of complete problems under our reduction technique for complexity classes arising from imposing upper bounds to the capabilities of local and global computations.

Eventually, it would be interesting to investigate properties that can(not) be decided by a frugal protocol with fixed number of rounds.

References

- [1] Noga Alon, Tali Kaufman, Michael Krivelevich, and Dana Ron. Testing triangle-freeness in general graphs. *SIAM J. Discrete Math.*, 22(2):786–819, 2008.

- [2] László Babai, Anna Gál, Peter G. Kimmel, and Satyanarayana V. Lokam. Communication complexity of simultaneous messages. *SIAM J. Comput.*, 33(1):137–166, 2004.
- [3] Yevgeniy Dodis and Sanjeev Khanna. Space time tradeoffs for graph properties. In *Proceedings of the 26th International Colloquium on Automata, Languages and Programming (ICALP)*, volume 1644 of *Lecture Notes in Computer Science*, pages 291–300. Springer, 1999.
- [4] Jan Draisma, Eyal Kushilevitz, and Enav Weinreb. Partition arguments in multiparty communication complexity. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming (ICALP)*, volume 5555 of *Lecture Notes in Computer Science*, pages 390–402. Springer, 2009.
- [5] Eldar Fischer. The art of uninformed decisions: A primer to property testing. *Bulletin of the EATCS*, 75:97–126, 2001.
- [6] Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, 1998.
- [7] Oded Goldreich and Dana Ron. Property testing in bounded degree graphs. *Algorithmica*, 32(2):302–343, 2002.
- [8] Stéphane Grumbach and Zhilin Wu. Logical locality entails frugal distributed computation over graphs (extended abstract). In *Proceedings of 35th International Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, volume 5911 of *Lecture Notes in Computer Science*, pages 154–165, 2009.
- [9] Daniel J. Kleitman and Kenneth J. Winston. On the number of graphs without 4-cycles. *Discrete Mathematics*, 41(2):167–172, 1982.
- [10] Alexandr V. Kostochka. Lower bound of the Hadwiger number of graphs by their average degree. *Combinatorica*, 4(4):307–316, 1984.
- [11] Fabian Kuhn, Thomas Moscibroda, and Roger Wattenhofer. What cannot be computed locally! In *Proceedings of the 23rd Annual ACM Symposium on Principles of Distributed Computing (PODC)*, pages 300–309. ACM, 2004.
- [12] Nathan Linial. Locality in distributed graph algorithms. *SIAM J. Comput.*, 21(1):193–201, 1992.
- [13] Peter Bro Miltersen. The bit probe complexity measure revisited. In *Proceedings of the 10th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, volume 665 of *Lecture Notes in Computer Science*, pages 662–671. Springer, 1993.
- [14] Peter Bro Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. On data structures and asymmetric communication complexity. In *Proceedings of the 27th Annual ACM Symposium on Theory of Computing (STOC)*, pages 103–111. ACM, 1995.
- [15] Mihai Patrascu and Erik D. Demaine. Logarithmic lower bounds in the cell-probe model. *SIAM J. Comput.*, 35(4):932–963, 2006.

- [16] David Peleg. *Distributed computing: a locality-sensitive approach*. SIAM Monographs on Discrete Mathematics and Applications, 2000.
- [17] Andrew Thomason. An extremal function for contractions of graphs. *Math. Proc. Cambridge Philos. Soc.*, 95:261–265, 1984.
- [18] Andrew Thomason. The extremal function for complete minors. *J. Comb. Theory, Ser. B*, 81(2):318–338, 2001.
- [19] Prasoos Tiwari. Lower bounds on communication complexity in distributed computer networks. *J. ACM*, 34(4):921–938, 1987.
- [20] E.M. Wright. Equal sums of like powers. *Bull. Amer. Math. Soc.*, 8:755–757, 1948.
- [21] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing (STOC)*, pages 209–213. ACM, 1979.
- [22] Andrew Chi-Chih Yao. Should tables be sorted? *J. ACM*, 28(3):615–628, 1981.